Reading to Learn: An Investigation into Language Understanding

Peter Clark[1], Phil Harrison[1], John Thompson[1], Rick Wojcik[1], Tom Jenkins[1], David Israel[2]
[1]Boeing Phantom Works, Seattle, WA
[2]SRI International, Menlo Park, CA
{peter.e.clark,philip.harrison,john.a.thompson,richard.h.wojcik,thomas.l.jenkins2}@boeing.com, israel@ai.sri.com

Abstract

One of the most important methods by which human beings learn is by reading. While in its full generality, the reading task is still too difficult a capability to be implemented in a computer, significant (if partial) approaches to the task are now feasible. Our goal in this project was to study issues and develop solutions for this task by working with a reduced version of the problem, namely working with text written in a simplified version of English (a Controlled Language) rather than full natural language. Our experience and results reveal that even this reduced version of the task is still challenging, and we have uncovered several major insights into this challenge. We describe our work and analysis, present a synthesis and evaluation of our work, and make several recommendations for future work in this area. Our conclusion is that ultimately, to bridge the "knowledge gap", a pipelined approach is inappropriate, and that to address the knowledge requirements for good language understanding an iterative (bootstrapped) approach is the most promising way forward.

Introduction

Overview

One of the most important methods by which human beings learn and increase their knowledge and understanding is by reading. Although in its full generality this task is still too difficult for computers, some significant, partial approaches are now feasible. Our goal in this research was to study one such reduced version of the problem, factoring the task as follows:

(i) "translate" the original English into a controlled language (a simplified version of English)
(ii) interpret that controlled language text

In this research, step (i) was performed manually, and step (ii) automatically. This allowed us to side-step some linguistic issues and instead concentrate on issues of machine understanding and knowledge integration. We are also able to speculate on the automatability of step (i).

While there has been substantial prior work in both controlled language processing (e.g., Fuchs et al, 1998; Mitamura et al., 2003) and generation of "logical forms" from full language (e.g., Moldovan and Rus, 2001), we seek to expand on this in two ways: First, our goal is to explore generating full, inference-supporting logic, capable of answering AP Chemistry exam questions. We thus have a high bar for what constitutes "machine understanding." Second, interpretation and inference is performed in the presence of an existing knowledge base, rather than a "blank slate", just as is the case for a human reading text.

Our work has been in the domain of chemistry, in which a prior, hand-built knowledge base already exists, created for Vulcan's Halo Pilot project (Barker et al, 2004; Friedland et al, 2004). Specifically, we have focused on 6 pages of chemistry text concerning acid-base equilibrium reactions, namely pp614-619 of (Brown, LeMay and Bursten, 2003). Our methodology was as follows:

1. Rewrite the 6 pages of chemistry text into our controlled language, CPL (Computer-Processable Language)
2. Use our CPL interpreter to generate logic from this
3. Integrate this new knowledge with an existing chemistry knowledge base (from the Halo Pilot)
4. Assess the performance of the CPL-extended KB on answering AP chemistry questions
5. Report problems encountered and solutions developed

Our experience and results, which we present here, reveal that even this reduced version of the task is still challenging, and we have uncovered several major insights into this challenge. In particular, our work reveals the need for an iterative (bootstrapped) approach to reading rather than a traditional "waterfall" approach, and for extensive use of background knowledge to guide interpretation. While this work has also involved substantial development of our CPL interpreter, we only mention features here where relevant; detail is provided in (Clark et al., 2005).

The Knowledge Gap

There is a fundamental gap between real natural language text, on one hand, and an "ideal" logical representation of that text that integrates easily with pre-existing knowledge, on the other. Importantly, this gap arises from more than just grammatical complexity; it involves multiple other factors that we describe in this paper. For full text comprehension, this gap must be bridged.

To reformulate the original text into CPL, there are two approaches, both of which we have investigated:

1. Write CPL which is "close" to the original English, i.e., is essentially a grammatical simplification of the original text with no/little new knowledge added. While this reformulation is done by hand, one can plausibly imagine this reformulation could be performed automatically by some suitable software.
2. Write CPL which fully captures the underlying knowledge that the author intended to convey, essentially treating CPL as a declarative rule language.

In this case, there is a significant gap between the original text and CPL reformulation, including significant new knowledge added to the reformulation.

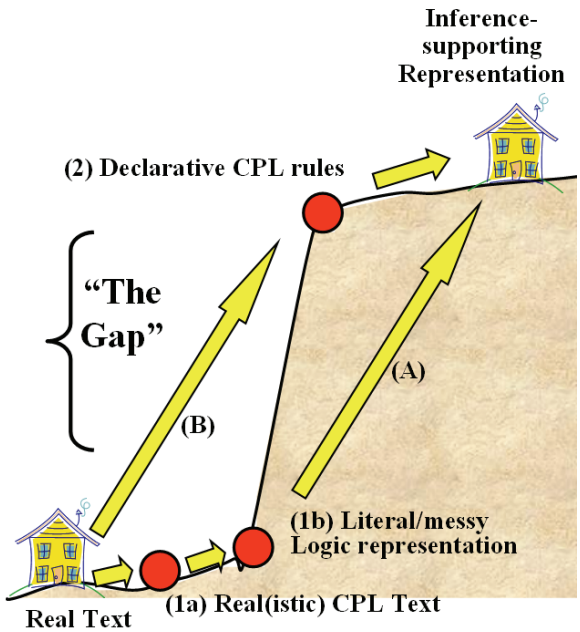


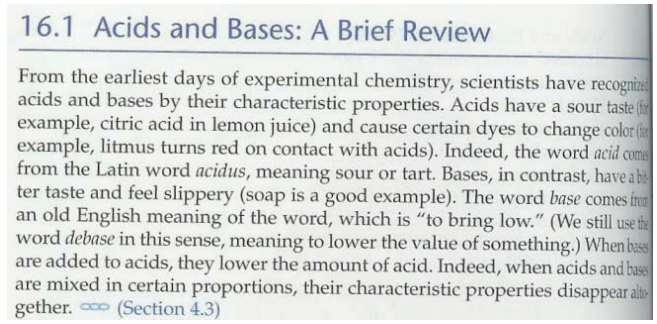
Figure 1: Paths between real text and useful logic

These two alternatives are shown in Figure 1, illustrating the "gap" between real language (the house at the bottom of the cliff) and inference-supporting logic (our goal, the house at the top of the cliff). The lower path illustrates the first alternative: While CPL close to the original text (1a) might plausibly be generated automatically, and logic generated from that (1b), a significant gap (A) still remains between that logic and that required to support inference. The upper path illustrates the second alternative: Writing the declarative rules underlying the text in CPL (2) crosses the gap (B), but only by virtue of significant and non-automatable manual intervention. In both formulations the "gap" between the messiness of real language, and the tidiness required for formal reasoning, is an obstacle. In this paper we provide a detailed analysis of this gap, its causes, and recommendations for how it might be bridged.

Analysis I: Sentence by Sentence Translation

Following the first (lower) path, in Figure 1, the six pages of chemistry text were rephrased as approximately 280 CPL sentences, and logic generated from them automatically by the CPL interpreter. Although some knowledge and selectivity was injected into these reformulations, they are still largely faithful to the original English, and the reformulation process might plausibly be automated. However, as we shall describe, the resulting logic is "messy" (bullet (1b) in Figure 1), essentially retaining much of the unwanted imprecision, overgenerality, and errors of the original text. These "problems" are things which a human reader typically does not even notice, and almost unconsciously he/she fills in and corrects the information he/she is reading. However, for a computer, they pose serious problems.

To illustrate this, consider the start of the text:



16.1 Acids and Bases: A Brief Review

From the earliest days of experimental chemistry, scientists have recognized acids and bases by their characteristic properties. Acids have a sour taste (for example, citric acid in lemon juice) and cause certain dyes to change color (for example, litmus turns red on contact with acids). Indeed, the word *acid* comes from the Latin word *acidus*, meaning sour or tart. Bases, in contrast, have a bitter taste and feel slippery (soap is a good example). The word *base* comes from an old English meaning of the word, which is "to bring low." (We still use the word *debase* in this sense, meaning to lower the value of something.) When bases are added to acids, they lower the amount of acid. Indeed, when acids and bases are mixed in certain proportions, their characteristic properties disappear altogether. (Section 4.3)

This text illustrates many typical challenges that arise. Consider the first sentence:

> *"From the earliest days of experimental chemistry, scientists have recognized acids and bases by their characteristic properties."*

To fully understand this requires already having some basic notion of time, chronologies, time periods, and their start and ends. It requires recognizing the idiom-like phrase "earliest days" as meaning "the start of". The sentence also includes generic references to scientists, acids, bases, and properties, and the challenge of interpreting generics (e.g., does it mean that all scientists recognize all acids all the time?). It includes a vague reference to "characteristic properties" - which properties, exactly, are being referred to there? Or how does this vague notion get recognized in the KB? Similarly, what sense of the verb "recognize" is intended here? This is particularly challenging as the author is not referring to specific recognition events, rather is referring to the state of understanding of scientists in the past and present. Later sentences in the paragraph similarly require prior knowledge about words and meaning.

In this particular paragraph, we have skipped much of this text as it is not central to the chemistry knowledge we are interested in. The CPL encoding we wrote looks:

```
Acids have a sour taste.
Acids cause some dyes to change color.
Bases have a bitter taste.
Bases have a slippery feel.
```

The logic generated from the first two sentences looks:

```
;;; Acids have a sour taste.
FORALL ?acid
  isa(?acid, Acid)
==> EXISTS ?taste
      isa(?taste, Taste-Value)
      taste(?acid, ?taste)
      value(?taste, *sour)
;;; Acids cause some dyes to change color.
FORALL ?acid
  isa(?acid, Acid)
```

```
==> EXISTS ?dye, ?color, ?change:
    isa(?color, Color-Value)
    isa(?dye, Substance)
    causes(?acid, ( isa(?change, Reaction)
            AND object(?change,?color)
            AND raw-material(?change,?dye))
```

Logic for the subsequent sentences follows a similar style. Technical details of the CPL interpreter are provided in (Clark et al., 2005).

In some cases, the generated logic is sensible, e.g., for "Acids have a sour taste", as the notions of taste and sour are known. In other cases the logic is not sensible, for a variety of reasons. An interesting case is the second sentence "Acids cause some dyes to change color." Taken literally, it is ambiguous, over-general, and erroneous:

• **Metonymy:** Strictly (at least in the Halo KB), only events can cause things, not objects. The sentence is referring to some (unstated) event involving acids that causes the change, and the word "acid" can be viewed as a metonymic reference to some event like "adding acid"

• **Presupposition**: The sentence omits (presupposes) contextual knowledge about how this change can take place, for example: The acid is in contact with the dye, the dye is not already the changed color, etc.

• **Ambiguity:** The sentence is ambiguous about whether the changing is a one-off or continuously ongoing event

• **Complex semantics:** The phrase "some dyes" really means "all instances of some types of dyes"; that is, it assumes prior knowledge that there is a natural grouping of dyes into types, and that each type is characterized by whether all its members change color with acids or not.

As a result, the logic representing the author's intended meaning would be substantially more complex and different than the "literal" logic produced by the CPL interpreter. Moreover, this logic would include additional knowledge not present in the original text (e.g., that the acid and dye must be touching); thus, it is in principle infeasible to generate this logic from the sentence alone - rather, substantial background knowledge is also needed (either preprogrammed or acquired through some learning process). Given all this, the logic that we have generated, taking a mechanical translation process which does not handle these issues, is largely unusable for meaningful inference. As we will show later in our second analysis, we can alternatively create richer CPL which avoids these problems, and generates inference-supporting logic - but of course we have then manually crossed the gap which we wish the machine to ultimately be able to bridge.

Synthesis

We have proceeded through the 6 pages of text in a similar manner, producing approximately 280 CPL sentences and corresponding logical clauses (Clark et al, 2006). Throughout the encoding process we encountered numerous encoding challenges which we have assembled into a catalog of 22 items, summarized below, along with a rough assessment of their degree of difficulty that they pose for a larger research program (green=easy, yellow=medium, red=difficult, black = extremely difficult). The example sentences below are all from the original chemistry text, with underlining added:

1. **Idioms/special-purpose phrases** (yellow)
   "*From the earliest days of experimental chemistry..."*
   Many phrases are meaningless when taken literally.
2. **Generics** (statements of general properties) (red)
   "Acids cause some dyes to change color."
   Challenges include unstated presuppositions, exceptions.
3. Handling **negation**. (green)
   "Some substances containing hydrogen are not acids"
4. **Vague attributes** (red)
   "their characteristic properties disappear altogether"
5. **coreference** (e.g., "react"/"reaction") (green)
   "*Hydrogen chloride reacts... The reaction produces.."*
6. **indirect anaphora**: reference to implied objects. (green)
   "Removing a proton from the acid produces the conjugate base."
7. acquiring **new technical vocabulary & meaning** (red)
   "NaOH dissociates in water."
   In particular, new meaning is challenging to acquire.
8. how to **represent definitions.** (green)
   "*An H-plus ion is a proton with no valence electron."*
9. how to state **generality** (green)
   "Bronsted-Lowry acids are more general than Arrhenius acids."
10. **modals/tendancies**, e.g., "can". (yellow/red)
    "A molecule of a Bronsted-Lowry acid can donate a proton..."
    A KR challenge, as no specific event is referenced.
11. how to represent **an argument (proof)**. (black)
    "Therefore, the H2O molecule acts as a Bronsted-Lowry base."
    Requires ability to represent and introspect on proofs.
12. **vagueness** ("is mostly", "nearby", "some") (red)
    "The NH4Cl is mostly solid particles."
    "Some acids are better proton donors than others"
    Ubiquitous even in chemistry, with complex semantics.
13. how to compute and represent **differences** (yellow)
    "An acid and a base differing only in a proton are called a conjugate pair"
14. **change over time** (yellow/red)
    "The HNO2 molecule becomes the NO2-minus ion."
    Requires modeling the dynamics of the world.
15. How to represent **hypothetical situations**. (yellow)
    "Assume that H2O is a stronger base than X-minus in Equation 16.9."
    Requires modeling alternative worlds.
16. **algebra:** how to reason with formulae and eqns (red)
    "OH- is the conjugate base of H2O"
    Chemistry is not just a world of chemicals, but also a world of symbols, and this second world also needs to be modeled. Terms like "OH-" are not just opaque identifiers but syntactic objects to be reasoned about.

17. **generalized formulae** and equations (black)
"In Equation 16.6 the symbol HX denotes an acid."
18. **loosespeak/metonymy** (yellow/red)
"*The H2O molecule in Equation 16 donates a proton"*
Ubiquitous in chemistry, eg. words for formulae, chemicals, and molecules are used interchangeably.
19. **Discourse context** (red)
"Every [Bronsted-Lowry] acid has a conjugate [Bronsted-Lowry] base"
Meaning is often dependent on previous sentences.
20. Descriptions **of problem-solving methods** (red)
"In any acid-base reaction we can identify two sets of conjugate acid-base pairs."
Reference to events not in the real world, but in the computational world of problem-solving.
21. **Generalization from examples** (red/black)
"We can identify two sets of conjugate acid-base pairs. For example, consider the reaction..."
22. Information **in tables and diagrams** (red/black)

Further quantification of these challenges is provided later.

Analysis II: Core Chemistry Knowledge

Introduction

In this second analysis, we look specifically for the knowledge required to answer AP-level questions, how it is expressed in the text, and how it can be encoded in CPL in an inference-capable way. Here, the CPL is substantially different from the original text, and CPL is essentially being used as an English-like rule language. This corresponds to taking the large leap (B) to cross the knowledge gap in Figure 1. One would hope that this "core" knowledge is explicitly and clearly presented in the original text, so that we can then transcribe it into CPL. However, this turns out not to be the case, and it is rarely stated in the nice explicit form that we would like.

Interestingly, the actual knowledge required for answering AP questions occupies only a small proportion of the text. For answering AP questions, there are 4 key competencies which need to be acquired:

1. Compute the conjugate base of an acid
2. Identify the strongest base from a pair of bases
3. Identify which are acids/bases in a reaction
4. Find direction of an equilibrium reaction (left/right)

We perform some detective work to find this knowledge in the original text, and look at how it can be manually expressed in inference-supporting CPL.

Case Study: Direction of Equilibrium

As a case study, consider the last competency, computing the direction of equilibrium. The context is an equilibrium reaction where acid1+base1 react to produce base2+acid2, the reaction simultaneously happening in reverse. Essentially, if base1 is "stronger" than base2, then equilibrium will be "to the right", i.e., acid1 + base1 will react more and as a result the equilibrium mixture will contain a greater proportion of base2+acid2 (the "right hand side" chemicals) than acid1+base1. This is explained in text mainly though two similar examples, the first being:

" $HX + H_2O \leftrightarrow H_3O^+ + X^-$.
If H_2O... is a stronger base than X^-...then H_2O will abstract the proton from HX to produce H_3O^+ and X^-. As a result, the equilibrium will lie to the right."

followed by a second example and statement of generality:

"From these examples, we conclude that in every acid-base reaction, the position of the equilibrium favors transfer of the proton to the stronger base."

This is all that is provided to explain this critical piece of knowledge. Although the above is (hopefully) sufficient to convey this notion to a person in context, it is formidable for a computer to process, even if it is rephrased into simplified English, and we encounter several of the challenges mentioned earlier. In particular:

1. **Generalization from examples** (challenge 21): The reader is largely meant to combine the examples with the rather vague, general statement to create a clear definition of equilibrium position. To write inference-capable CPL, we have to perform this combination manually.
2. **Special phrases** (challenge 1): "lies [to the right]" is a somewhat metaphorical phrase simply meaning "is"; similarly "favors transfer" is a complex, metaphorical phrase indicating greater concentration. This must be made explicit in the CPL.
3. **Metonymy** (challenge 18): The sentences mix the chemical, molecular, and symbolic worlds together. For example, in the last general statement "transfer of the proton to the stronger base" really means "transfer of the proton to a molecule in the stronger base". A "literal" interpretation of the original text would be incorrect, and as a result extra knowledge needs to be inserted in the CPL to make the meaning explicit. Similarly, the notions of "left" and "right" refer to the equation, not the physical world.

In this second analysis we have manually untangled and clarified the knowledge. The resulting CPL rule looks:

```
IF there is an equation of a reaction
AND a 1st chemical has a formula
AND a 2nd chemical has a second formula
AND the 1st formula is part of the left side
of the equation
AND the 2nd formula is part of the right
side of the equation
AND the 1st chemical is playing a base role
AND the 2nd chemical is playing a base role
AND the 1st chemical is stronger than the
2nd chemical
THEN the direction of the reaction is right
AND the equilibrium side of the reaction is
right.
```

Note that much of the condition part of this rule is dealing with metonymy in the original text, i.e., it is explicitly converting from a reaction, to its equation, to the formulas in it, to the chemicals they denote. Also the metaphorical references have become explicit general principles.

The Other Three Competencies

We might hope that the indirectness of the original text for this last competency is unusual. Unfortunately, this is not the case: the English describing the other three core tasks is similarly obtuse, at least to the computer, also relying on diagrammatic information and assuming general problem-solving know-how by the reader (Clark et al., 2006).

Thus, although we can author CPL rules which will successfully chain together to answer AP questions, they are substantially different to the original text, and highly unlikely to be derived automatically from the text. The "knowledge gap" thus still remains a formidable problem.

Knowledge Integration and Extensible KBs

Our goal with our CPL-generated knowledge is that it is integratable with the existing chemistry KB. As part of our experimental methodology, our goals included surgically removing the hand-written acid-base knowledge from the original Halo KB, adding in the CPL generated knowledge, and comparing performance.

Ideally, our CPL generated knowledge would look similar to the equivalent hand-built knowledge in the KB, so we could just remove the latter and insert the former. Unfortunately, this is not the case; the original hand-built knowledge is highly complex and intertwined. In this section we briefly study the previous encoding of this chemistry knowledge, look at why it is complex and why this makes knowledge integration hard, and then reflect on how we would like that original knowledge to have looked so that knowledge integration would be easier.

As an example, consider the 2nd competency, namely identifying the strongest base from a pair of bases. In the original KB, rather than represent a partial ordering on strength, the KB authors decided to use three absolute, qualitative, strength values: “strong”, “weak”, and “negligible”, and then encode the comparison operator as a large, nested if-then structure specifically for computing acid strengths:

<u>RULE (English paraphrase of original in Halo KB):</u>
IF acid X is strong & acid Y is not strong
THEN X is a stronger acid than Y
  ELSE IF acid X is not strong & acid Y is strong
  THEN Y is a stronger acid than X
    ELSE IF acid X is weak & acid Y is negligible
    THEN X is a stronger acid than Y
      ELSE IF acid X is negligible &acid Y is weak
      THEN Y is a stronger acid than X.

There are several challenges this representation poses:

1. The encoding is not easy to automatically extend, due to its **syntactic complexity**; a knowledge integration algorithm would have to tease apart and reassemble this structure with any desired modifications. A preferable encoding would be using ground assertions, e.g., a table.

2. The representation **mixes general knowledge** about reasoning with ordered scales **and specific knowledge** about acids. It would be preferable to separately represent knowledge about reasoning with ordered scales, and then apply that knowledge in the specific context of reasoning about acid strengths. Clancey has made similar observations about encoding knowledge in expert systems (Clancey, 1984; Clancey, 1992)

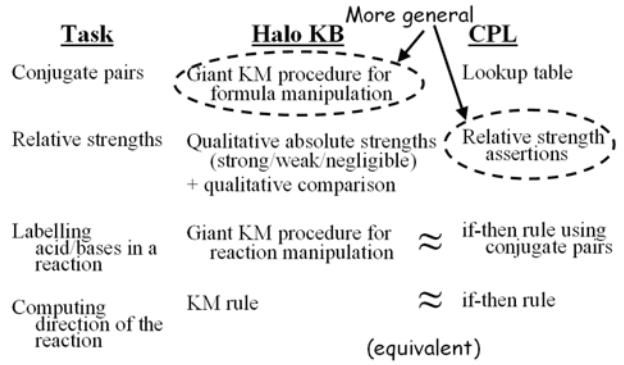
3. This representation **hard-wires a 3-value scale** for acid strengths. A preferable encoding would be to have general knowledge about ordered scales generalized to N-values.

Again, one might ask if this is just an “unlucky example”. However, investigation of the encoding of the other competencies reveal similar challenges (Clark et al, 2006). The bottom line is that extensibility needs to be deliberately designed into the representation (this was not a goal for the original KB). Thus learning by reading does not just require good linguistic processing, it also demands careful design of the target representational structures if the read knowledge is to be easily integrated into them.

Evaluation and Quantification of the Gap

Performance on AP Questions

Our original goal was to compare the performances of the CPL-generated knowledge with that of the original Halo KB on AP Chemistry questions. In fact, because of the relatively short amount of text we are dealing with, we can predict the performance on any given question by analysis, by assessing the generality of the encoded methods (looking at the Analysis II CPL rules):



For example, the Halo KB can answer more questions than CPL for Task 1 (conjugate pairs), but less for Task 2. The net result of this analysis is that, for any given sample of AP questions, we can identify by analysis which KB will score highest. While we could in principle perform this experiment for real, this would tell us little beyond the statistical distribution of question types in the AP exam, something that is tangential to our concerns here. Rather, the most significant thing we have learned from our work

is the size and nature of the "knowledge gap" to be bridged. As a result, we have instead focused our concluding analysis on a rough quantification of this gap.

A Quantification of the Knowledge Gap

In the first analysis we listed a number of phenomena contributing to the knowledge gap. In this Section, we aggregate these into some broader categories, and present a rough quantification of their prevalence in the 6 pages of chemistry text that we have studied. For comparison, we also compare this with their prevalence in six pages of grade-school level biology text about a different subject (the structure and function of the heart), to give some indication of which phenomena are domain general and which are accentuated in college-level chemistry. The 22 categories earlier were aggregated as follows:

- Idioms/special-purpose phrases (item 1 earlier)
- Generics (item 2)
- Knowledge representation challenges (items 3-15)
- Algebra/mathematics (items 16-17)
- Loosespeak/metonymy (item 18)
- Discourse context (item 19)
- Problem-solving method descriptions (item 20)
- Learning from examples (item 21)
- Tables and diagrams (item 22)

The prevalance of these phenomena was identified by counting the numbers of sentences in which they occur. Note this is a very loose quantification given the relatively small amount of text (six pages) looked at in each science.

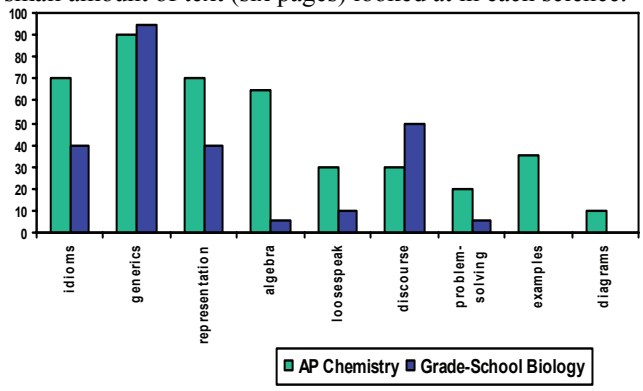


Table 1: Quantification of the relative frequency (% of sentences) of the nine major challenges, as seen in our target chemistry text and a comparably sized biology text.

There are some interesting, albeit tentative, conclusions which can be drawn from this data. First, AP level chemistry text is considerably harder to interpret computationally than grade-school level biology text. Part of this stems from the educational level of the text: grammatically, the biology text contained shorter, simpler sentence structures, which were more likely to stand on their own than the chemistry text. Use of idiomatic phrases/phrases with idiosyncratic meaning was less common in the biology text. Semantically too, the complexity of the knowledge being communicated is higher at the AP level, resulting in a higher number of "representational challenges" to address than for grade-school level text. As a second dimension of comparison, the discipline makes a difference also: Biology is a subject concerned with structure, function, and their relationship in real-world settings - these are things which AI systems can model relatively well. In contrast, chemistry, at least at the AP level, includes reasoning at the molecular as well as real-world level, and the use of equations and formulae is central to modeling what is happening. All these pose extra challenges for machine understanding of text. Thus, it appears that some of the complexities of chemistry are not universal, but rather peculiar to that particular science (or class of sciences), and that machine understanding may prove less challenging in other domains.

Summary and Conclusions

We have conducted an investigation into the opportunities and challenges of having a machine read to learn, focusing on part of the domain of chemistry, and simplifying the challenge by working with controlled language. Our work shows that while grammatical simplification of the language helps, significant challenges still remain. We have identified a catalog of linguistic challenges. In addition, our analysis illustrates that the final KB must be deliberately designed for extensibility if automatic knowledge integration is to be achieved.

Despite the variety of these issues, they essentially all require the use of extensive, prior knowledge to help guide and correct the interpretation. This creates a kind of "catch 22": reading can add knowledge, but only if there is reliable knowledge there to begin with. If a system had sufficient prior knowledge about plausible relationships between objects in a domain, for instance, it would then be able to recognize and correct metonymy or other imprecisions in the input language. How can a system acquire such knowledge in the first place? While some hand-coding may be feasible, it seems clear that only practical way forward is through some bootstrapped/ looping approach, in which some initial knowledge supports at least some reading, which then augments knowledge, supporting further reading etc. In terms of our original metaphor of the knowledge gap, rather than thinking of reading as a task of "climbing a cliff", it is perhaps better thought of as a two way process, where knowledge flows down from the top, providing expectations, context, prior knowledge, and hypotheses for interpreting language, and new information from language provides fragments of new knowledge, examples, confirmations, and refinements. Thus rather than thinking of language processing as a pipeline, viewing it in terms of a symbiotic relationship between language and knowledge is perhaps a more appropriate road to success. While some early NLP work has emphasized expectations (e.g., Schank and Abelson, 1977), there have been few examples of bootstrapped systems that use newly acquired knowledge to guide future reading. This work reinforces that need, and we are hopeful that such an approach will lead to improved technologies for machine reading in the future.

Acknowledgements

This work was funded under DARPA contract HR0011-05-C-0073.

References

K. Barker, V. Chaudhri, S. Chaw, P. Clark, J. Fan, D. Israel, S. Mishra, B. Porter, P. Romero, D. Tecuci, P. Yeh. A Question-Answering System for AP Chemistry: Assessing KR&R Technologies. In Proc 9th International Conf on Knowledge Representation and Reasoning (KR'04), 2004, AAAI Press.

T. Brown, E. LeMay, H. Bursten. Chemistry: The Central Science. NJ:Prentice-Hall, 2003.

W. Clancey. Classification Problem Solving. In AAAI'84, pp 49-55, 1984.

W. Clancey. Model Construction Operators. In Artificial Intelligence 53 (1), pp1-116, 1992.

P. Clark, P. Harrison, T. Jenkins, J. Thompson, R. Wojcik. Acquiring and Using World Knowledge using a Restricted Subset of English. In: The 18th International FLAIRS Conference (FLAIRS'05), 2005.

P. Clark, P. Harrison, T. Jenkins, J. Thompson, R. Wojcik, D. Israel. Reading to Learn – Final Report. Available at http://www.cs.utexas.edu/users/pclark/rtol/ 2006.

J. Fan, K. Barker, B. Porter. Indirect Anaphora Resolution as Semantic Search. Proc. 3rd International Conference on Knowledge Capture, pp153-160, 2005.

D. Fass Met*: A Method for Discriminating Metonymy and Metaphor by Computer, in Computational Linguistics 17 (1), 49-90, 1991.

N. Friedland, P. Allen, M. Witbrock, G. Matthews, N. Salay, P. Miraglia, J. Angele, S. Staab, D. Israel, V. Chaudhri, B. Porter, K. Barker, P. Clark. Towards a Quantitative, Platform-Independent Analysis of Knowledge Systems. In Proc 9th International Conf on Knowledge Representation and Reasoning (KR'04), 2004, AAAI Press.

N. E. Fuchs, U. Schwertel, R. Schwitter. Attempto Controlled English - Not Just Another Logic Specification Language. In: Proc. 8th Workshop on Logic-Based Program Synthesis and Transformation (LOPSTR'98), 1998.

T. Mitamura, K. Baker, E. Nyberg, D. Svoboda. Diagnostics for Interactive Controlled Language Checking. In: Proc. Controlled Language Application Workshop (CLAW'03), 2003.

D. Moldovan, V. Rus. "Logic Form Transformation of wordNet and its Applicability to Question Answering". In: Proceedings of the ACL 2001 Conference, July 2001, Toulouse, France.

R. Schank, R. Abelson. Scripts, Plans, Goals and Understanding. NJ: Erlbaum. 1977.